CROPPING SYSTEMS

Evaluation of GPFARM for Dryland Cropping Systems in Eastern Colorado

Allan A. Andales, Lajpat R. Ahuja,* and Gary A. Peterson

ABSTRACT

GPFARM is an ARS decision support system for strategic (long-term) planning. This study evaluated its performance for comparing alternative dryland no-till cropping systems and established limits of accuracy for eastern Colorado, using data collected in 1987 through 1999 from an ongoing long-term experiment at three locations along a gradient of potential evapotranspiration (PET) (Sterling, low PET; Stratton, medium PET; and Walsh, high PET). The crop rotations, which included winter wheat (*Triticum aestivum* L.), corn (*Zea mays* L.), sorghum [*Sorghum bicolor* (L.) Moench], proso millet (*Panicum miliaceum* L.), and varying fallow periods, were wheat–fallow, wheat–corn–fallow, and wheat–corn–millet–fallow at Sterling and Stratton and wheat–fallow, wheat–sorghum–fallow, and wheat–sorghum–millet–fallow at Walsh. The ranges of relative error (RE) of simulated mean and root mean square error (RMSE) were total soil profile water content (RE: 0 to 23%; RMSE: 38 to 76 mm water), dry mass grain yield (RE: −27 to 84%; RMSE: 419 to 2567 kg ha^{-1}), dry mass crop residue (RE: −5 to 42%; RMSE: 859 to 1845 kg ha^{-1}), and total soil profile residual nitrate N (RE: −42 to 32%; RMSE: 26 to 78 kg ha^{-1}). GPFARM simulations agreed with observed trends and showed that productivity and water use efficiency increased with cropping intensification and that Stratton was the most productive and Walsh the least. GPFARM (v. 2.01) was less suited for year-to-year grain yield prediction under dryland conditions but has potential as a tool for studying long-term interactions between environment and crop management system. Future development and applications of GPFARM must account for crop-specific responses to stress, detailed hydrology, better understanding of root uptake processes, and spatial variability to give more accurate grain yield predictions in water-stressed environments.

Sustainable agriculture demands consideration of many interrelated factors, processes, resources, and institutions. In the Great Plains, there has been a recognized need for a systems approach for agricultural research and development for attaining sustainability (Ascough et al., 2002). Peterson et al. (1993) proposed that a systems approach to the study of soil and crop management problems is useful for testing present research knowledge to answer practical agricultural problems and simultaneously identify gaps in basic research knowledge. Likewise, there has been a recognized need for system-level decision support tools for agricultural advisors and producers. In a 1995 Great Plains survey of 121 county extension directors, 173 NRCS district conservationists, and 95 agricultural consultants, more than 90% were interested in using farm management decision support software (Frasier et al., 1997). The same survey also showed that 57% of 219 producer respondents were interested in a farm management decision support product. Central to meeting this challenge, the USDA-ARS Great Plains Systems Research Unit (GPSR), in a collaborative effort with Colorado State University, developed the GPFARM (Great Plains Framework for Agricultural Resource Management) decision support system (DSS) (Ascough et al., 2002; Shaffer et al., 2000). The general purpose of GPFARM is to serve as a whole farm/ranch DSS in strategic planning across the Great Plains, for production, economic and environmental impact analysis, and site-specific database generation, from which alternative agricultural management systems can be tested and compared. The term strategic planning is defined here as long-term planning (e.g., choice of sustainable crop rotation, choice of tillage/residue management system, etc.) as opposed to tactical planning (e.g., scheduling of irrigation, chemical application, harvesting, etc.), which is done in real time. Agricultural consultants and progressive farmers or ranchers are targeted as the primary users of GPFARM. The user requirements for GPFARM were identified by ARS customer focus groups in the Great Plains, comprised of farmers, ranchers, agricultural consultants, and NRCS and extension professionals. The major requirements were that (i) the DSS be simple to understand and easy to use and (ii) have minimum input data and parameter requirements.

The GPFARM model is an aggregate of modules taken from existing agricultural water quality models and new modules specifically developed for GPFARM. For example, the crop growth module is based on the EPIC generic crop growth model that has been widely tested for various crops (e.g., Steiner et al., 1987; Williams et al., 1989; Martin et al., 1993; Moulin and Beckie, 1993; Kiniry et al., 1995; Jara and Stockle, 1999) while the water balance module, which is a simplification of the RZWQM (Ahuja et al., 2000) water balance routines, has not been extensively tested.

Most of the modules have been independently tested to varying degrees, but there is a need to evaluate GPFARM at the system level to see how well the mod-

A.A. Andales and L.R. Ahuja, USDA-ARS Great Plains Syst. Res. Unit, P.O. Box E, Fort Collins, CO 80522; and G.A. Peterson, Dep. of Soil and Crop Sci., Colorado State Univ., Fort Collins, CO 80523. Received 29 Aug. 2002. *Corresponding author (Laj.Ahuja@ars.usda.gov).

Published in Agron. J. 95:1510–1524 (2003).

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: CCC, Colorado Climate Center; CV, coefficient of variation; DSS, decision support system; ET, evapotranspiration; HI, harvest index; LAI_{max}, maximum leaf area index; PET, potential evapotranspiration; RE, relative error of mean; RMSE, root mean square error; WCF, wheat–corn–fallow (rotation); WCMF, wheat–corn–millet–fallow (rotation); WC(S)F, wheat–corn (or sorghum)–fallow (rotation); WC(S)MF, wheat–corn (or sorghum)–millet–fallow (rotation); WF, wheat–fallow (rotation); WSF, wheat–sorghum–fallow (rotation); WSMF, wheat–sorghum–millet–fallow (rotation); WUE, water use efficiency.

ules work together to simulate various cropping systems, especially for conditions in the immediate target area of eastern Colorado. GPFARM is currently being evaluated in five ways: (i) on-farm/ranch testing, (ii) research plot or scientific testing, (iii) expert opinion evaluation by producers and scientists, (iv) sensitivity analysis, and (v) trend analysis (McMaster et al., 2003). Deer-Ascough et al. (1998) made preliminary evaluations of the grain yield simulations of the generic crop growth module in GPFARM for winter wheat, corn, and proso millet. They tested the GPFARM model for dryland wheat–fallow (WF), wheat–corn–fallow (WCF), and wheat–corn–millet–fallow (WCMF) rotations in eastern Colorado. The average relative grain yield prediction error across three sites in eastern Colorado and three cropping rotations was 30%. Since this preliminary evaluation, many corrections and enhancements have been made to the GPFARM science modules. Therefore, the main objectives of this study were (i) to evaluate the overall performance of GPFARM version 2.01 in simulating alternative dryland cropping systems in eastern Colorado over multiple years (long term), (ii) to identify limits of reliability within which GPFARM can be used as a strategic planning tool, and (iii) to assess the value of such simpler modeling approaches in practical applications for the future. A secondary objective was to identify the limitations of the model that warrant further investigation, more rigorous testing of specific modules, reparameterization, or reworking of theories and mechanisms, especially if it were also to be used for year-to-year planning.

MATERIALS AND METHODS

The GPFARM Decision Support System

The GPFARM DSS is unique in that it brings together a suite of decision support tools integrated with a complex whole farm/ranch simulation model and relational databases that are accessible through a user-friendly interface that was designed specifically for producers through close collaboration with several cooperators in the Great Plains. The main contribution of GPFARM is not the introduction of new science but rather the delivery of current research knowledge, embodied in the simulation model and built-in databases, to agricultural producers and advisors in a user-friendly form. For ease of development and reduction of parameters, the developers used simpler scientific approaches that hopefully would be adequate in distinguishing alternate management systems for long-term strategic planning. Databases of model input parameters based on the literature were integrated into the DSS. Parameterization of plant, soil, climate, and other components are performed for the user, and all other inputs are minimized as much as possible (McMaster et al., 2003). Therefore, the GPFARM simulation model is a compromise between scientific rigor and simplicity.

The GPFARM DSS is an aggregate of six major components designed to serve as an extensive decision support tool for farmers and ranchers (Fig. 1). The first component is a Microsoft Windows–based graphical user interface (GUI) that facilitates the entry of input data, provides simulation control, and displays output results. The second component includes Microsoft Access databases containing the soils, crops, weeds, climates, agricultural implements, chemicals, and economic parameters needed in the simulations and analysis of results. The third component is an object-oriented modeling framework that integrates modules for simulating soil water dynamics, N dynamics, crop growth, weed growth, beef cattle production, pesticide transport, and water/wind erosion. The fourth component is a set of analysis tools including a multicriteria decision-making model (MCDM), graphical/spatial output visualization, and summary report tables—all to help analyze and compare different management scenarios. The fifth component is a stand-alone economic analysis tool that can take production data either from the science model or from user input to perform detailed economic analyses on the farm or ranch enterprise. The sixth component is the Internet-based

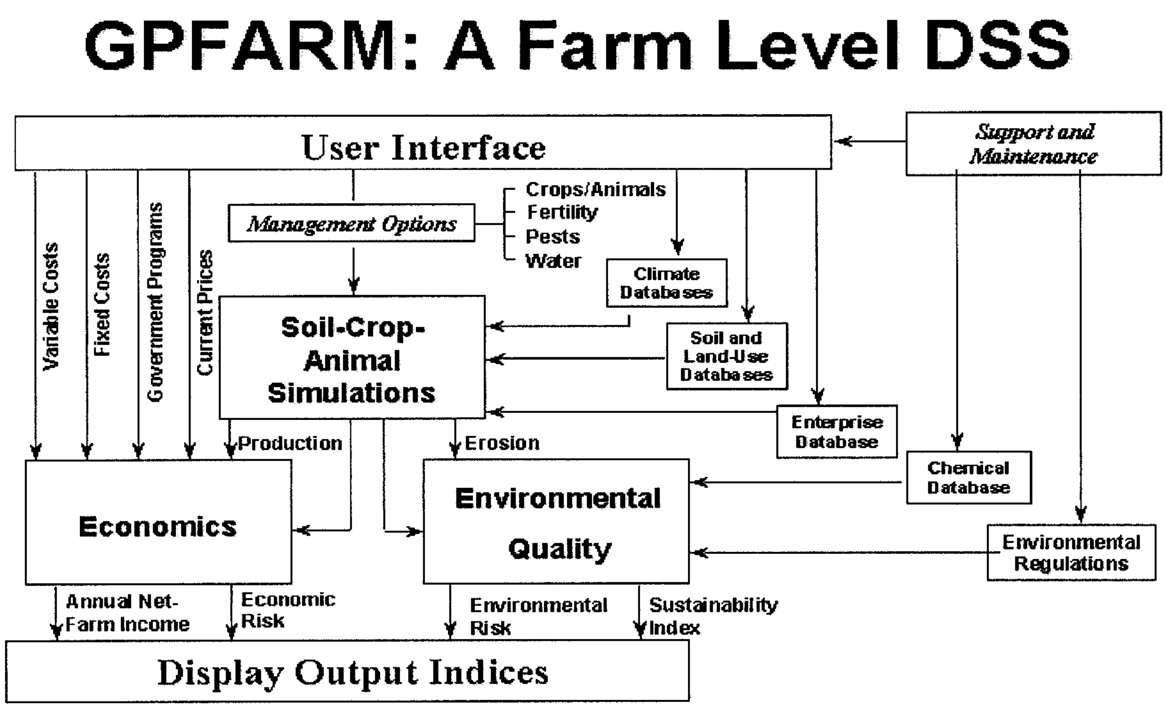

Fig. 1. Schematic diagram of the GPFARM decision support system (DSS) components. Arrows indicate the flow of information.

(http://infosys.ars.usda.gov/; verified 1 Aug. 2003) GPFARM information system containing numerous links to information on various farm and ranch management options. The system contains information on crops and crop management; range and pasture management; livestock production; soil, water, and nutrient management; and weed/pest control.

The remainder of this section will be limited to an overview of the science simulation modules that were run to produce the simulation results presented in this paper. Ascough et al. (2002) and McMaster et al. (2003) present more comprehensive overviews of the GPFARM DSS.

The GPFARM science model is a field-by-field simulation framework (Shaffer et al., 2000) using object-oriented programming in C++ and executes appropriate simulation modules, written in procedural languages (FORTRAN and BASIC), for the basic processes (e.g., crop growth, water dynamics, and C and N cycling). The pertinent simulation modules for this study were the crop growth module, the soil properties module, the PET module, the water balance and chemical transport module, and the C- and N-cycling module.

- Crop growth module. Deer-Ascough et al. (1998) describe this module, which is a modified version of the EPIC crop growth submodel (Williams et al., 1989) as used in the WEPP soil erosion model (Arnold et al., 1995). The module uses concepts of daily accumulated heat units for plant phenology; Monteith's approach for determining potential biomass (Monteith, 1977); simple conceptual water, N, and temperature stress adjustments to daily growth; and harvest index (HI) for partitioning biomass to economic yield. Crop/variety-specific parameters to simulate daily growth are kept in a default database. Currently, GPFARM is parameterized for winter wheat, corn, sunflower (*Helianthus annuus* L.), sorghum, proso millet, and foxtail/hay millet [*Setaria italica* (L.) Beauv.].

- Soil properties module. This module estimates the soil water retention curve based on the Brooks and Corey (1964) parameters from soil texture, bulk density, and organic matter content (Rawls and Brakensiek, 1985). The above soil property information is obtained from the soil survey database or provided by the user. The saturated hydraulic conductivity is obtained from effective porosity (Ahuja et al., 1989). Unsaturated hydraulic conductivity is estimated from the water retention curve and saturated hydraulic conductivity using the Campbell (1974) approach. The effects of tillage, residue cover, and reconsolidation (due to rainfall) on bulk density are estimated using the approach of Williams et al. (1984), and hydraulic properties are updated using the regression equations of Rawls and Brakensiek (1985).

- PET module. This module, which was adapted from the RZWQM (Ahuja et al., 2000), calculates daily potential crop transpiration and soil evaporation using the extended Shuttleworth–Wallace model (Farahani and Ahuja, 1996). The module calculates net radiation and partitions the available energy for potential transpiration, bare soil evaporation, and/or residue-covered soil evaporation. The potential transpiration, potential soil evaporation, and potential residue evaporation values then serve as the upper limits of actual evapotranspiration (ET) calculated in the water balance module.

- Water balance and chemical transport module. This module is a simplification of the RZWQM water balance routines (Ahuja et al., 2000) and uses a coarser time step (RZWQM time step ranges from 10^{-5} to 1 h; GPFARM time step ranges from 1 h to 6 h) between precipitation events to determine soil water fluxes. Water supply at the surface comes from natural precipitation, irrigation, or snowmelt. A simple disaggregation scheme is used to convert the daily rainfall inputs to intensities of an average daily rainstorm to simulate infiltration and runoff. The Green–Ampt (Green and Ampt, 1911) method is used to simulate infiltration during a rainstorm at small time intervals while redistribution of soil water is by Darcian flux (Darcy, 1856) calculated at 3-h to daily intervals between adjacent layers. Surface water supply exceeding the infiltration capacity in any time interval of precipitation becomes surface runoff. Soil evaporation is a function of soil water content of the first 5 cm of the top soil layer and is limited by the Darcy flux toward the surface and the potential soil evaporation. Actual transpiration is the sum of root water uptake from each soil layer, which is based on root mass distribution in the profile, available water, and the potential transpiration. Drainage from the soil profile is estimated by assuming a unit gradient at the bottom layer. Chemical transport is coupled with water movement based on a uniform mixing model and partitioned between aqueous and adsorbed fractions. Pesticide degradation is simulated as a first-order process with a known half-life.

- C- and N-cycling module. Based on the NLEAP model (Shaffer et al., 1991, 2001), this module simulates soil C and N cycling in surface residues and within the soil. It has two soil organic matter pools (a fast, readily decomposable pool and a slower humus pool) and one surface residue pool (Shaffer et al., 2001). Each pool has its own C/N ratio and is subject to first-order decomposition. The processes of nitrification, ammonia volatilization, denitrification, crop N uptake, and nitrate N leaching are also included.

Dryland Agroecosystem Management Project

Experimental data for model testing were taken from an ongoing, pioneering, dryland agroecosystem project in eastern Colorado (Peterson et al., 1993). The overall objective of the project is to identify dryland crop and soil management systems that maximize plant water use efficiency (WUE) and maintain soil productivity while providing an economically sustainable level of production. Peterson et al. (1993) give a detailed description of the experimental design, and only an abbreviated version of their description is given here. The experimental design is a split block that includes climatic environment (low PET, medium PET, and high PET), slope position (summit, sideslope, toeslope), and cropping system variables. The climatic environment variable was represented by the three locations, all in eastern Colorado, representing three levels of PET: Sterling (low ET; 40.37° N, 103.13° W), Stratton (medium ET; 39.18° N, 102.26° W), and Walsh (high ET; 37.23° N, 102.17° W). Long-term average annual precipitation values are 440, 415, and 395 mm for Sterling, Stratton, and Walsh, respectively. Long-term cropping season open pan evaporation averages 1600, 1725, and 1975 mm for Sterling, Stratton, and Walsh, respectively. Each location was divided into two blocks (replicates). Within each block, three cropping systems were present: WF, wheat–corn (or sorghum for the Walsh site)–fallow [WC(S)F], and wheat–corn (or sorghum for the Walsh site)–millet–fallow [WC(S)MF]. Each phase of each cropping system (rotation) was randomly assigned within each block. Thus, at each location, all phases of each rotation were present in two replications each year. An experimental unit is a particular phase of a cropping system at a particular slope

Table 1. Soil physical and hydraulic properties at Sterling, Stratton, and Walsh, CO.

Soil horizon	Bulk density	Sand	Clay	OM†	Porosity‡	WC§ (33 kPa)‡	WC (1500 kPa)‡	Ksat‡¶
cm	g cm^{-3}	%			m^3 m^{-3}			cm h^{-1}
				Sterling summit [Weld loam (fine, smectitic, mesic Aridic Paleustoll)]				
0–8	1.37	45.0	20.8	1.37	0.48	0.24	0.14	3.31
8–20	1.35	33.4	30.8	1.09	0.49	0.30	0.17	1.20
20–30	1.23	24.5	38.1	1.09	0.52	0.34	0.20	0.86
30–51	1.21	27.4	27.0	0.77	0.54	0.31	0.17	2.70
51–69	1.31	30.5	22.1	0.46	0.51	0.28	0.14	2.48
69–85	1.34	43.4	20.7	0.21	0.49	0.25	0.14	3.67
85–120	1.43	31.1	25.4	0.13	0.46	0.29	0.15	0.86
				Stratton summit [Norka clay loam (fine-silty, mixed, mesic Aridic Argiustoll)]				
0–13	1.41	25.00	34.00	1.76	0.47	0.33	0.19	0.23
13–39	1.34	20.00	36.00	1.48	0.49	0.35	0.21	0.27
39–43	1.31	29.00	25.00	0.77	0.51	0.29	0.14	2.27
43–64	1.31	27.00	21.00	0.46	0.51	0.29	0.14	2.27
64–96	1.37	35.00	18.00	0.28	0.48	0.25	0.13	2.86
96–150	1.35	34.00	14.00	0.00	0.49	0.24	0.11	3.95
				Walsh summit [loamy sand (fine-loamy, mixed, mesic, Aridic, Ustochrept)]				
0–18	1.48	65.40	14.30	0.60	0.44	0.18	0.11	5.22
18–40	1.49	66.40	17.50	0.40	0.44	0.18	0.11	4.27
40–68	1.46	67.60	16.70	0.28	0.45	0.18	0.11	5.13
68–96	1.41	61.00	19.20	0.26	0.47	0.21	0.13	4.74
96–135	1.39	56.40	21.10	0.26	0.48	0.22	0.13	4.10
135–150	1.17	6.50	39.90	1.02	0.56	0.39	0.22	0.49

† OM, organic matter.
‡ Estimated in GPFARM.
§ WC, water content.
¶ Ksat, saturated hydraulic conductivity.

position within a block at a site. The size of an individual experimental unit varies. All units are 6.1 m wide but vary in length with the particular site (ranging from 185–305 m). However, a constant length in the middle of each unit was harvested for the experiment.

Cropping systems represent a continuum with increasing cropping intensity and fewer summer fallow periods. The cropping systems were all managed with no-till techniques to maximize water storage potential. Since all phases of each rotation were present each year, all cropping systems could be compared on an annual basis because all crops in a given system were annually present and were affected by that year's particular environmental conditions.

Dramatic differences in soils exist between the three sites. Table 1 shows the measured soil physical properties along with the soil hydraulic properties estimated by GPFARM. The summit loam soil at Sterling is relatively shallow, with a partially cemented layer at about 90-cm depth that is slowly permeable to water but relatively impermeable to roots. At Stratton, the summit soil is clay loam with few water or root restrictions. The summit soil at Walsh is loamy sand with no restrictions to water infiltration or root penetration and a *plug* at 135-cm depth by virtue of the abrupt increase in clay content. In the order of decreasing plant available water-holding capacity, the summit soils are ranked Stratton > Walsh > Sterling.

Fertilizer N was applied to each experimental unit, according to soil tests obtained from each soil within each rotation and specific for the crop present in a given year. The N fertilizer source, urea NH_4NO_3 solution (32–0–0), was applied at planting with a dribble method directly behind the planter (Peterson et al., 1993). Phosphorus (10–34–0) was band-applied at planting of all crops near the seed (Peterson et al., 2000). Phosphorus was applied on one-half of each wheat (until 1992), corn, and proso millet plot over all soils but applied to the entire wheat plot since the 1993 crop year. The P application rate was 9.5 kg/ha each year. Grain yield response to P is not currently simulated in GPFARM, but it was found to be small to negligible in the experiment.

Measurements pertinent to the evaluation of GPFARM included daily weather data, soil water content, soil residual NO_3–N, dry matter biomass and grain yields, and crop residue dry mass. Many more variables were measured, as described by Peterson et al. (1993), but were not considered in the model evaluations. An automated weather station at each site measured daily air temperature (maximum and minimum), mean relative humidity, precipitation, total solar radiation, wind direction, and mean wind speed. Soil water content (30-cm increments down to a depth of 150 cm) was measured at strategic times (biweekly during summer months) in each cropping system by use of neutron attenuation. Soil residual NO_3–N (at varying increments down to a depth of 150 cm) was measured before planting for making fertilizer N calculations. Dry mass grain yields were measured with a plot combine while total aboveground biomass was measured at harvest by hand-sampling a small area in each experimental unit. The harvest indices (dry mass grain yield/total biomass) were determined from the hand samples. Crop residue dry mass was measured at planting and just before harvest for each crop in each cropping system. With the exception of weather variables, all measurements from a particular cropping system were done in two replicates (i.e., taken from two blocks). The replicates were averaged for comparison with simulation results from GPFARM.

Model Inputs and Calibration

Model Inputs

The GPFARM model was initialized using observed data for soil profile water content, crop residue, and soil profile residual nitrate N corresponding to the simulation start dates. Observed bulk density, texture, and organic matter content of the soil layers (Table 1) were also input into GPFARM. From these properties, the model estimated the soil water retention curve, soil porosity (or saturated water content), soil water content at field capacity (33 kPa), soil water content at wilting point (1500 kPa), and saturated/unsaturated hydraulic

Table 2. Best estimates of critical crop parameters used in GPFARM simulations.

Parameter	Definition	Units	Value			
			Winter wheat	Corn	Proso millet	Sorghum
GDDMAX	growing degree days from planting to maturity/harvest	°C-d	2300.00	1500.00	1300.00	1800.00
HI†	harvest index	0–1 ratio	0.48	0.65	0.45	0.50
HMAX	maximum canopy height	m	0.91	2.60	1.20	1.01
LAI_{max}†	maximum leaf area index potential	$m^2\ m^{-2}$	2.00	3.50	2.40	3.50
BEINP	biomass energy ratio for a crop	$kg\ MJ^{-1}$	30.00	35.00	35.00	25.00
BN1	plant N concentration parameter at seedling stage	$kg\ t^{-1}$	0.0600	0.0400	0.0440	0.0440
BN2	plant N concentration parameter halfway through the season	$kg\ t^{-1}$	0.0231	0.0164	0.0164	0.0164
BN3	plant N concentration parameter at maturity	$kg\ t^{-1}$	0.0134	0.0128	0.0128	0.0128
BTEMP	base temperature of crop	°C	0.00	10.00	5.00	10.00
CRIT	growing degree days to emergence	°C-d	140.00	60.00	65.00	60.00
DLAI	heat unit index when leaf area index starts to decline; fraction of GDDMAX	0–1	0.70	0.80	0.80	0.85
EXTNCT	radiation extinction coefficient		0.65	0.65	0.65	0.60
OTEMP	optimal temperature for plant growth	°C-d	20.00	25.00	20.00	27.50
RDMAX	maximum rooting depth	m	1.50	1.50	1.00	1.50
RSR	root/shoot ratio	0–1 ratio	0.25	0.25	0.25	0.25
SPRIOD	period over which senescence occurs	d	14.00	30.00	30.00	40.00
RLAD	rate of LAI decline		1.00	0.10	1.00	1.00
PPOP1	plant density at FMLAI1	$plants\ m^{-2}$	125.00	4.00	125.00	5.00
FMLAI1	fraction of XMXLAI corresponding to PPOP1	0–1 ratio	0.60	0.47	0.60	0.43
PPOP2	plant density at FMLAI2	$plants\ m^{-2}$	250.00	7.00	250.00	15.00
FMLAI2	fraction of XMXLAI corresponding to PPOP2	0–1 ratio	0.95	0.80	0.80	0.79

† Calibrated to optimize predicted total biomass, HI, and grain yield.

conductivity (Rawls and Brakensiek, 1985; Ahuja et al., 1989, 1999). Actual soil horizons (Table 1) were used for the simulations. Actual N application rates were also used in the simulations.

Weather inputs into GPFARM included observed daily precipitation (mm), maximum and minimum air temperatures (°C), solar radiation (Langleys d^{-1}), wind speed ($m\ s^{-1}$), and relative humidity (%). All weather inputs were measured onsite, except for daily precipitation data for Sterling, which were downloaded from the Colorado Climate Center (CCC) website. The CCC precipitation station near Sterling was approximately 21 km (13 miles) northwest of the experimental site. The complete precipitation data set (1987–1999 for Sterling) from the CCC was used in lieu of the on-site precipitation data set that had numerous gaps, especially during the winter months.

Model Calibration

Model calibration was done only for two plant growth parameters: the maximum leaf area index (LAI_{max}) and the potential HI. For other plant growth parameters of the crops involved in the study (winter wheat, corn, proso millet, and sorghum), the best estimates from the literature were used (Table 2) and verified to be within the ranges recommended by Arnold et al. (1995) and Kiniry et al. (1995). For soil water, soil residual nitrate N, and crop residue decomposition processes, no calibrations were done.

The LAI_{max} for each crop was adjusted (within ranges expected for the study site) to minimize the RMSE of simulated total aboveground biomass. The HI for each crop was adjusted by trial and error (based on observed HI) to minimize the RMSE of HI predictions. Input values for HI represent potential (unstressed) values. The calibrated HI values for winter wheat (HI = 0.48) and corn (HI = 0.65) were considerably higher than those recommended by Kiniry et al. (1995), which were 0.40 and 0.55, respectively. Nevertheless, the simulated harvest-time HI values, which were adjusted for water, temperature, and N stresses, ended up much lower and were close to observed values. Apparently, the high calibrated HI values for winter wheat and corn compensated for overadjustments of HI in the model.

The input values of LAI_{max} and HI are supposed to be values for nonstress conditions with respect to water, temperature, or N (Williams et al., 1989). These unstressed input values are then adjusted inside the model for stresses. In an adjunct study, we compared the use of summit vs. toeslope data from the Sterling site to obtain these calibrated values for winter wheat, corn, and proso millet. Data from the WF rotation (1988–1997) beginning with the wheat phase in 1988 were used for winter wheat calibration, data from the WCF rotation (1988–1999) beginning with the corn phase in 1988 were used for corn calibration, and data from the WCMF rotation (1988–1993) beginning with the millet phase in 1988 were used for proso millet calibration. The summit soil profile at Sterling had a root restriction at 90-cm depth and had less available water than the toeslope position, with a deeper soil profile and more total available water. The calibrated values of LAI_{max} and HI obtained from the above two data sets were significantly different, obviously affected by different degrees of stress (data not shown). It also turned out that the parameters calibrated based on the small subset of experimental data at the summit position gave better predictions of grain yields at the summit for the rest of the summit data than the parameters calibrated based on toeslope data. In spite of these results, we chose to use the parameters calibrated from the toeslope data at Sterling for model evaluations because they represented much lower stress conditions, closer to the nonstress conditions theoretically required, and because their use provided a more rigorous (independent) test of the crop model at the summit positions of the three sites.

In the calibration at the Sterling toeslope position, simulated grain yield agreed with observed values, and simulated total soil profile water content was slightly lower than observed at most times (Fig. 2). The LAI_{max} and HI values for sorghum were calibrated using Walsh data (WSF rotation beginning with the sorghum phase in 1988) because sorghum was planted only at that location.

Model Evaluation Procedure

To address the main objectives of the study, we focused on answering the following sets of questions:

1. How accurate are GPFARM simulations of total soil profile water content, grain yield, crop residue, and total residual soil profile nitrate N?
2. Can GPFARM simulate cropping system differences in

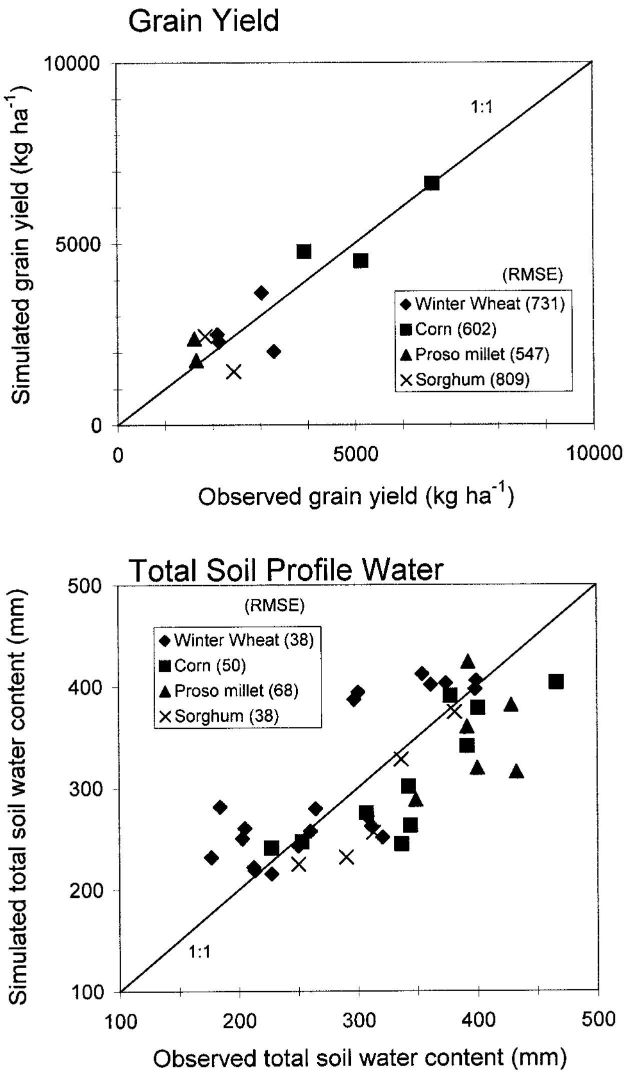

Fig. 2. Simulated grain yield and total soil profile water content against the observed values for calibration years at the toeslope position of the Sterling site. The simulated soil profile was 150 cm deep.

productivity within and across locations over multiple years?

3. What are the strengths and limitations of the model?

For answering the above questions, we chose to limit our evaluations to summit positions of the Sterling, Stratton, and Walsh locations because the sideslope and toeslope positions had the uncertainty of receiving unmeasured runoff water from upslope areas, which is not simulated in the present version of GPFARM. The evaluations of soil water and soil residual nitrate N simulations were limited to total soil profile amounts.

As this study comprised the first system-level test of GPFARM, our crop model evaluations addressed the ability to simulate weather-induced (i.e., caused by water and temperature stresses) variability in grain yields at each location. We felt that this should first be established before attempting to simulate other factors (e.g., natural hazards) that affect dryland grain yields in eastern Colorado where grain yields are usually limited by precipitation. Observed grain yields that were affected by weed infestation, erratic emergence due to hard surface soil conditions, hail damage, or killing frost were excluded from comparisons with simulated grain yields. Steiner et al. (1987) and Cabelguenne et al. (1999) used similar approaches of data screening to limit evaluations to the validity domain of the models. The aforementioned adverse factors, which are not unusual in the Great Plains, will need to be added and tested separately in the future for GPFARM to be widely applicable in the Great Plains. GPFARM includes a weed module, but there were insufficient quantitative observations of weed infestation to allow calibration of the weed module. At present, we are not aware of any single crop model that can adequately simulate all of the adverse factors mentioned.

The simulation periods for evaluation began in 1988 and ended in 1997, 1999, and 1993 for WF, WC(S)F, and WC(S)MF rotations, respectively. The WF simulations ended in 1997 because this system was subsequently converted to wheat–corn–proso millet rotation (Peterson et al., 2000). For WC(S)MF, sunflower was planted in place of proso millet after 1993 (Peterson et al., 1995), but the sunflower crops produced little or no yields, which limited our ability to calibrate the crop model for sunflower grain yield. Thus, we ended the WC(S)MF simulations in 1993. Average (i.e., from two replicates) total soil profile water content, grain yield, crop residue, and total residual soil profile nitrate N observed during the above periods were compared with corresponding GPFARM simulation outputs. In the calculation of evaluation statistics, we pooled data from all phases of a rotation at each location. For example, using the WCF rotation at Sterling, observed (mean of two replicates) and simulated data from the wheat phase (WCF-W), corn phase (WCF-C), and fallow phase (WCF-F) were pooled to calculate each statistic describing the WCF rotation at Sterling.

The following four statistics were calculated to quantify the accuracy of the GPFARM simulations: (i) RE, which shows bias of the predicted mean relative to the observed mean; (ii) RMSE, which shows the average deviation between predicted and observed values, regardless of sign; (iii) index of agreement, d, which gives the proportion of the observed variance that is explained by the model; and (iv) simulated and observed coefficient of variation, CV, which show whether or not simulated and observed variability are similar. Simulated values were compared with the mean of two observations (replicates) from each cropping system. Relative error was expressed in percent as:

$$\mathrm{RE} = \frac{(\bar{p} - \bar{o})}{\bar{o}} 100 \qquad [1]$$

where $\bar{p}$ is the predicted mean and $\bar{o}$ is the observed mean. The RMSE was calculated by:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (p_i - o_i)^2}{n}} \qquad [2]$$

where p_i is the ith predicted value, o_i is the ith observed value, and n is the number of data pairs. The index of agreement was calculated as proposed by Willmott (1981) and Willmott and Wicks (1980):

$$d = 1 - \left[\frac{\sum_{i=1}^{n} (p_i - o_i)^2}{\sum_{i=1}^{n} (|p_i'| + |o_i'|)^2}\right], \; 0 \le d \le 1 \qquad [3]$$

where p_i, o_i, and n are as previously defined and $p_i' = p_i -$

$\bar{o}$ and $o_i' = o_i - \bar{o}$, where $\bar{o}$ is the observed mean and the enclosing slashes (//) indicate absolute values. A d value of 1 indicates complete agreement between model predictions and observations.

Differences in overall productivity between the WF, WC(S)F, and WC(S)MF rotations were assessed through comparison of annualized yields from the rotations for the period 1989 through 1993. [WC(S)MF rotation only existed in the experiment until 1993.] We calculated annualized yield by summing the dry mass of grain yields of all crops and dividing by the total number of years in the rotation. Simulated and observed annualized yields were compared to check for reasonable simulation of trends in productivity associated with increased cropping intensity. Two sets of comparisons were made to consider: (i) the effect of increasing cropping intensity and (ii) the effect of climate and soil (i.e., differences in locations).

Water use efficiency, production per unit water used, is a diagnostic tool for evaluating cropping systems with a single numeric value because it combines productivity and water use (Peterson et al., 1993). The grain yield WUE diagnostic is an important and sensitive means of evaluating the combined effects of climate, soil, and cropping system. Water use efficiency was calculated for the period 1989 through 1993 by dividing total dry mass grain yield by the total ET for the entire period. Note that the fallow periods were included. Simulated and observed WUE were compared to evaluate the ability of GPFARM to distinguish between the performances of different cropping systems.

Water use efficiency for each cropping system in a given period was calculated using the equation:

$$\text{WUE} = \text{grain yield/ET} \qquad [4]$$

where grain yield is the total dry mass grain yield for the period (kg ha^{-1}) and ET is the total evapotranspiration for the period (cm H_2O). We had some difficulty in estimating ET from the experiment because on-site measurements of some water balance components (i.e., surface runoff and drainage) were not available. Thus, our calculation of *observed* ET was limited by availability of measured data, and we had to settle on the following approximation for both simulated and observed ET from each cropping system:

$$\text{ET} = \text{WC}_i + \text{precip} - \text{WC}_f \qquad [5]$$

where WC_i is initial soil water content in the profile (cm), precip is the total precipitation during the period (cm), and WC_f is the final soil water content in the profile at the end of the period (cm). Equation [5] assumes that surface runoff and drainage from the profile are negligible, which is not a bad assumption in the long term for this semiarid region and was also assumed by Peterson et al. (1993) in their WUE calculations.

RESULTS

Evaluation of Process Simulations

The GPFARM model simulations of dryland cropping systems at three locations (summit positions) in eastern Colorado were evaluated based on four process state variables: total soil profile water content, grain yield, crop residue, and total soil profile residual NO_3–N. Results of the quantitative evaluation are described below for each state variable.

Total Soil Profile Water Content

Overall, total soil profile water content simulations were better at Walsh (Fig. 3; lowest RMSE and highest d values) than at Sterling or Stratton. This may be attributed to better soil parameterization and precipitation data at Walsh. The REs in the means ranged from 13 to 23% at Sterling, 18 to 23% at Stratton, and 0 to 11% at Walsh. Mean water contents at the three locations were generally overpredicted, but the tendency for overprediction was more evident at Sterling and Stratton (Fig. 3). The reasons for the overprediction could not be identified because of lack of experimental information on some water balance components (e.g., surface runoff) and on root distribution in the soil profile. Simulated and observed variability in soil water content, rep-

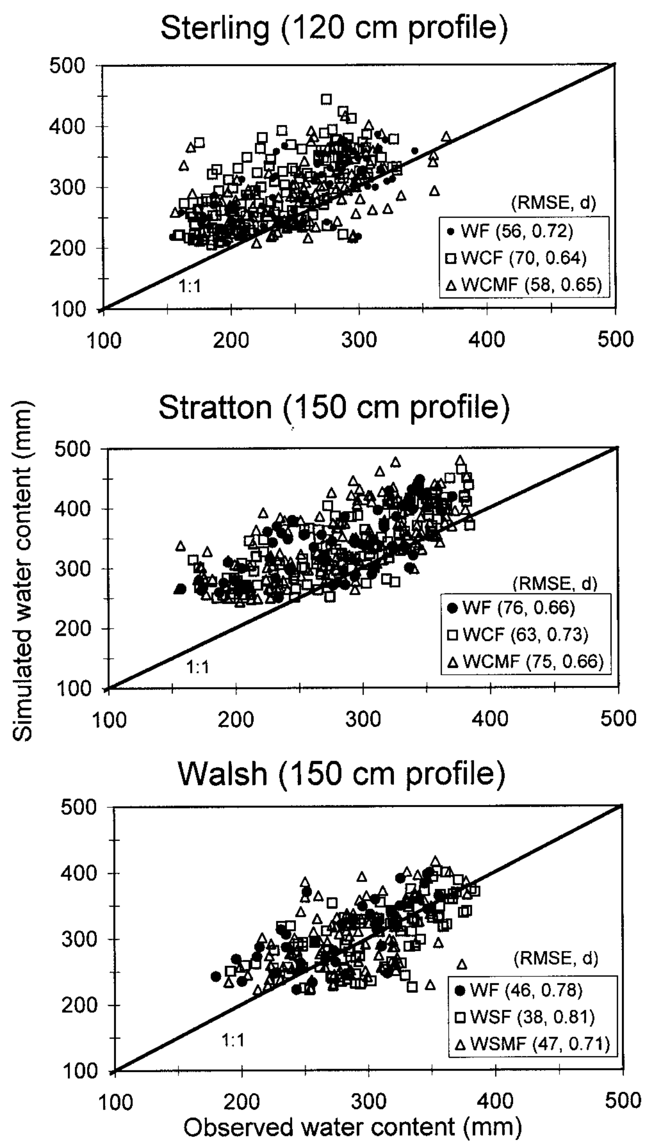

Fig. 3. Simulated total soil profile water content against the observed values at three locations (summit position) in eastern Colorado for three rotations {wheat–fallow (WF): 1988–1999; wheat–corn (or sorghum)–fallow [WC(S)F]: 1988–1999; wheat–corn (or sorghum)–millet–fallow [WC(S)MF]: 1988–1993}. Values in parentheses are root mean square error (RMSE) (mm H_2O) and d values.

resented by the CVs, were similar at Sterling and Walsh, and simulated variability was slightly less than observed at Stratton (detailed statistics not shown).

GPFARM simulated the correct timing for most of the observed drying and wetting events over time for all locations. In the case of the WC(S)F rotation beginning with the corn (or sorghum) phase in 1988, the simulated soil profile did not dry out as much as observed at Sterling and Stratton, but closely followed the drying and wetting patterns at Walsh (Fig. 4).

Grain Yield

A comparison of RE values among the Sterling, Stratton, and Walsh locations reveals that GPFARM generally predicted long-term average grain yields with a margin of ±30% or better (Table 3). Exceptions were corn yield prediction at Sterling (≈50% RE) and millet yield prediction at Walsh (84% RE). Based on RMSE values, the best simulations of grain yield were for winter wheat, and the worst were for corn (see Fig. 5 and 6). Among the four crops, the lowest RE values and highest *d* values were obtained with winter wheat. Apparently, as a long-duration crop, winter wheat is less sensitive to variable soil moisture conditions than the short-duration summer crops (corn, sorghum, and millet).

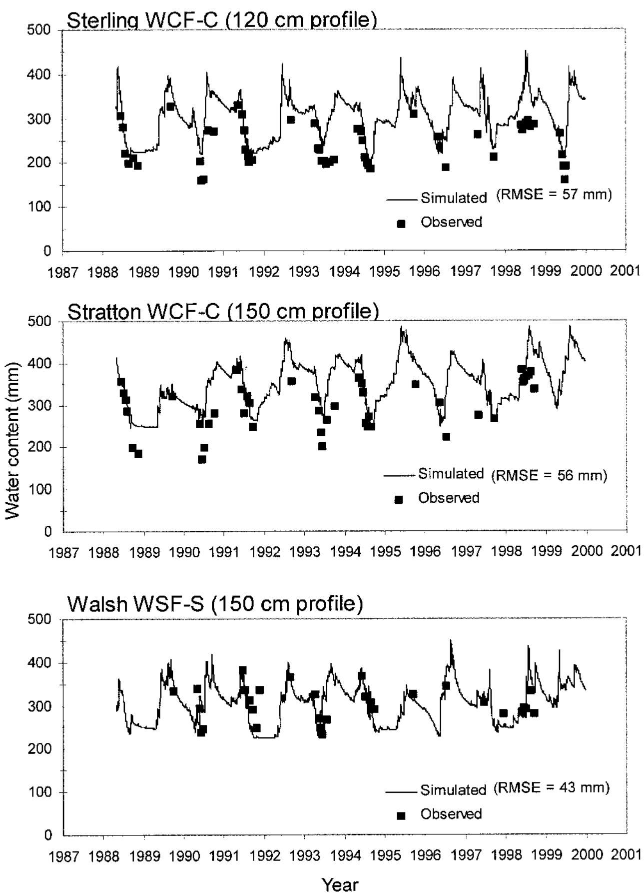

Fig. 4. Time series (1988–1999) of simulated and observed total soil water content at Sterling, Stratton, and Walsh for the wheat–corn (or sorghum)–fallow [WC(S)F] rotation beginning with the corn (or sorghum) phase in 1988 (Sterling and Stratton: WCF-C; Walsh: WSF-S).

Grain yield prediction for winter wheat and proso millet tended to be better at the location where they were calibrated. This was not the case for corn. For instance, REs for winter wheat were generally lower (and *d* generally higher) at Sterling than at Stratton and Walsh. The RE of predicted millet yield was lower (RE = 11.5%) at Sterling than at Stratton (RE = 27%) or Walsh (RE = 84%).

The REs in simulated winter wheat grain yield were within ±27%, with the lowest magnitudes occurring at Sterling (where winter wheat was calibrated) and the largest magnitudes at Walsh (Table 3). Mean winter wheat grain yields were generally overestimated at Sterling and underestimated at Stratton and Walsh. The RMSE values for winter wheat grain yield were generally lower at Sterling (where winter wheat was calibrated at the toeslope position) than at Stratton and Walsh (Fig. 5). The *d* values for winter wheat grain yield varied over a wide range (0.26–0.78). Oddly, both the highest (WCMF: $d = 0.78$) and lowest (WCF: $d = 0.26$) *d* values were observed at Sterling. In general, there was a tendency to underestimate wheat grain yield variability at Sterling and to overestimate variability at Stratton and Walsh. Simulated CV tended to decrease with increasing mean simulated winter wheat grain yields (Table 3).

The REs in simulated corn grain yield were larger than for winter wheat grain yield. Mean corn grain yields were overpredicted by around 50% at Sterling and by 29% at Stratton. The RMSE values for corn grain yield (Fig. 6a and 6b) were the highest (>2000 kg/ha) among the four crops. They were slightly lower at Sterling (where corn was calibrated at the toeslope position) than at Stratton. Agreement between simulated and observed corn grain yields was poor (d = 0.16–0.31) at Sterling and mediocre ($d \approx 0.50$) at Stratton. The simulated CV was similar to observed in all cases except for WCMF at Sterling.

For sorghum, which was planted only at Walsh, the RE was approximately 0% in the WSMF rotation and −23% in the WSF rotation (Table 3). The RMSE values were 686 kg/ha in the WSMF rotation and almost double (1016 kg/ha) in the WSF rotation (Fig. 6c). The *d* values were 0.77 and 0.53 for the aforementioned rotations, respectively. The CV of sorghum grain yield was overestimated.

Relative errors of proso millet grain yield simulations were 12% at Sterling and 27% at Stratton (Table 3). The evaluation statistics at Walsh were not very meaningful because there were only two observations. Because of consistently low yields, proso millet at Walsh was replaced by forage sorghum beginning in 1993. Similar to our observations for winter wheat and corn, the RMSE for proso millet grain yields was lower at Sterling (where millet was calibrated at the toeslope position) than at Stratton and Walsh (Fig. 6d). There was poor agreement between simulated and observed proso millet grain yields at Sterling and Stratton (d = 0.18–0.44). The CV was underestimated at Sterling and overestimated at Stratton.

Overall, the grain yield simulations leave much to be desired and indicate limitations of the crop model in

Table 3. Evaluation statistics for simulated grain yield (three locations in eastern Colorado; three rotations).

Rotation†	No. of observations	Observed Mean	Observed CV‡	GPFARM simulated Mean	GPFARM simulated CV	RE§	d (0–1)
		kg ha^{-1}	%	kg ha^{-1}	%	%	
Winter wheat grain yield							
Sterling							
WF	7	1929	27	1918	27	−1	0.65
WCF	10	1969	33	2216	16	13	0.26
WCMF	5	1964	31	2165	20	10	0.78
Stratton							
WF	8	2459	20	1887	44	−23	0.35
WCF	8	2180	16	2173	25	0	0.44
WCMF	9	2463	23	2148	31	−13	0.50
Walsh							
WF	6	1702	30	1296	26	−24	0.58
WSF	6	1971	23	1433	30	−27	0.66
WSMF	6	1868	14	1648	32	−12	0.41
Corn grain yield							
Sterling							
WCF	8	3315	25	5043	29	52	0.31
WCMF	6	3024	16	4464	37	48	0.16
Stratton							
WCF	5	3780	47	4877	47	29	0.49
WCMF	5	3723	43	4795	46	29	0.50
Sorghum grain yield (Walsh only)							
WSF	5	2577	24	1989	46	−23	0.53
WSMF	11	1995	36	1998	42	0	0.77
Proso millet grain yield [WC(S)MF only]							
Sterling	5	1810	22	2018	12	12	0.18
Stratton	5	1554	28	1975	47	27	0.44
Walsh	2	814	1	1496	35	84	0.00

† WF, wheat–fallow; WCF, wheat–corn–fallow; WCMF, wheat–corn–millet–fallow; WSF, wheat–sorghum–fallow; WSMF, wheat–sorghum–millet–fallow.
‡ CV, coefficient of variation.
§ RE, relative error (of mean).

simulating grain yield under dryland conditions in eastern Colorado. We looked into the biomass and HI simulations at Sterling to get some insight into the shortcomings of the crop model. In the model, grain yield is calculated by multiplying the aboveground biomass at harvest by the HI (adjusted for water, temperature, or N stress). We found that mean corn biomass was overpredicted by 35 to 45% whereas mean simulated HIs were similar to observed values. For winter wheat, the RE values were within ±10% for both biomass and HI, but the agreement between simulated and observed HI was poor to mediocre (d = 0.22–0.51). The variability in winter wheat biomass was also underestimated (i.e., simulated CVs lower than observed). For proso millet, there was poor agreement between simulated and observed biomass (d = 0.20) and HI (d = 0.24) while the RE values were still within ±15%. Variability in both biomass and HI of proso millet was also underestimated. Errors in prediction of biomass seem to be the major reason for errors in simulated grain yield for corn and proso millet whereas in winter wheat, the contributions of biomass and HI to errors in simulated grain yields varied with rotation.

Crop Residue

In the majority of cases, the RE in residue prediction was less than 26% (detailed statistics not shown). The RE values were lower (−5% to 16%) at Sterling and Stratton than at Walsh (16–42%). The RMSE values (Fig. 7a) were generally lowest at Sterling and tended to increase with cropping intensity at all locations [i.e., WC(S)MF exhibited the highest RMSE values compared with WF and WC(S)F]. Also, better agreement between simulated and observed crop residue was obtained at Sterling and Stratton (d = 0.67–0.80) than at Walsh (d = 0.45–0.59). The CV was generally underestimated at all locations, with the simulated CV being closer to the observed CV at Sterling than at Stratton or Walsh. Amounts of surface crop residue are closely tied to amounts of crop biomass produced. The model assumes that 80% of stalks are added to existing surface crop residue at harvest. Thus, errors in biomass prediction translate to errors in crop residue prediction. Inaccuracies in the simulation of residue addition during harvesting and subsequent decay may have also contributed to errors.

Total Soil Profile Residual Nitrate Nitrogen

The prediction of residual soil profile nitrate N amounts at planting time is inherently complex because of numerous plant–soil–environment factors that interact to influence N cycling in the soil. Predicting nitrate N amounts over an extended number of years is an even greater challenge. Overall, the soil residual nitrate N was predicted within ±40% RE or better. The RE values for total soil profile residual NO_3–N varied widely at each location (detailed statistics not shown). The RE was exceptionally low under WF at Sterling (−2%) and under WCF (−1%) and WCMF (5%) at Stratton. The means were consistently underestimated at Walsh (RE = −42% to −20%). The RMSE values were lowest at Sterling and greatest at Stratton (Fig. 7b). The highest d

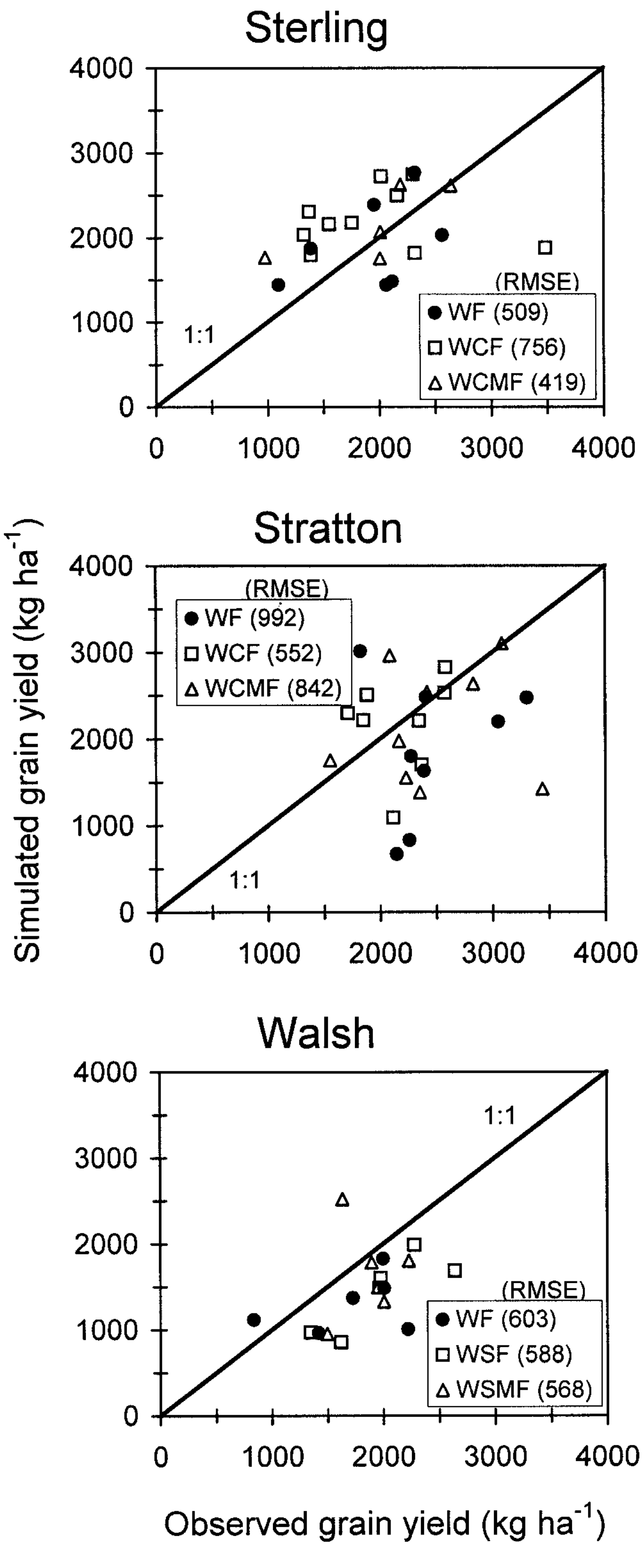

Fig. 5. Simulated winter wheat grain yield (dry mass) against the observed values at three locations (summit position) in eastern Colorado for three rotations {wheat–fallow (WF): 1988–1999; wheat–corn (or sorghum)–fallow [WC(S)F]: 1988–1999; wheat–corn (or sorghum)–millet–fallow [WC(S)MF]: 1988–1993}. Values in parentheses are root mean square error (RMSE) values (kg ha^{-1}).

values were obtained at Sterling (WF: d = 0.82; WCF: d = 0.60) while d values in all other cases ranged from 0.25 to 0.59. There was a tendency to overpredict soil residual nitrate N variability in most cases. Predicted residual nitrate N was highly sensitive to the amount of organic matter in the soil and to crop leaf area index. The lack of within-season residual nitrate N data prevented us from evaluating root uptake of nitrate.

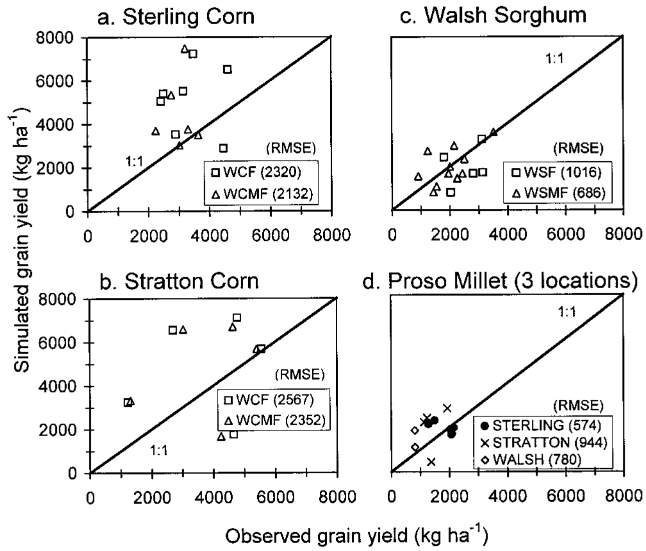

Fig. 6. Simulated dry mass corn grain yield at (a) Sterling and (b) Stratton, simulated sorghum grain yield at (c) Walsh, and (d) simulated proso millet grain yield {wheat–corn (or sorghum)–millet–fallow [WC(S)MF] rotation only} at three locations against the observed values for two rotations {wheat–corn (or sorghum)–fallow [WC(S)F]: 1988–1999; WC(S)MF: 1988–1993}. Values in parentheses are root mean square error (RMSE) values (kg ha^{-1}).

Evaluation of Cropping System Simulations

The model was able to simulate two observed trends among the cropping systems (Peterson et al., 1993). First, the model simulated the increased productivity with a 3- or 4-yr rotation vs. the 2-yr rotation (Fig. 8). Second, the model simulated productivity differences between locations: Stratton being the most productive and Walsh being the least productive. Annualized simulated grain yields followed a similar trend as the observed, but deviations from the observations were more pronounced for the WC(S)F and WC(S)MF systems. This was expected as the corn and proso millet grain yields were overpredicted for those systems (Table 3). Annualized yield predictions for the WF system were very close to the observed at all locations. The results demonstrate that, although it did not capture the spatiotemporal variability in grain yield very well (see Table 3), GPFARM may be useful in evaluating cropping systems on the basis of long-term relative productivity. Future improvements in the accuracy of grain yield simulations will increase the accuracy of annualized grain yield comparisons among alternative cropping systems.

Simulated and observed grain WUE for the three cropping systems at the three locations showed similar trends (Fig. 9). The simulated and observed grain WUE for each cropping system may be inaccurate because of the assumptions (i.e., surface runoff and drainage were negligible) made in calculating ET, but the differences

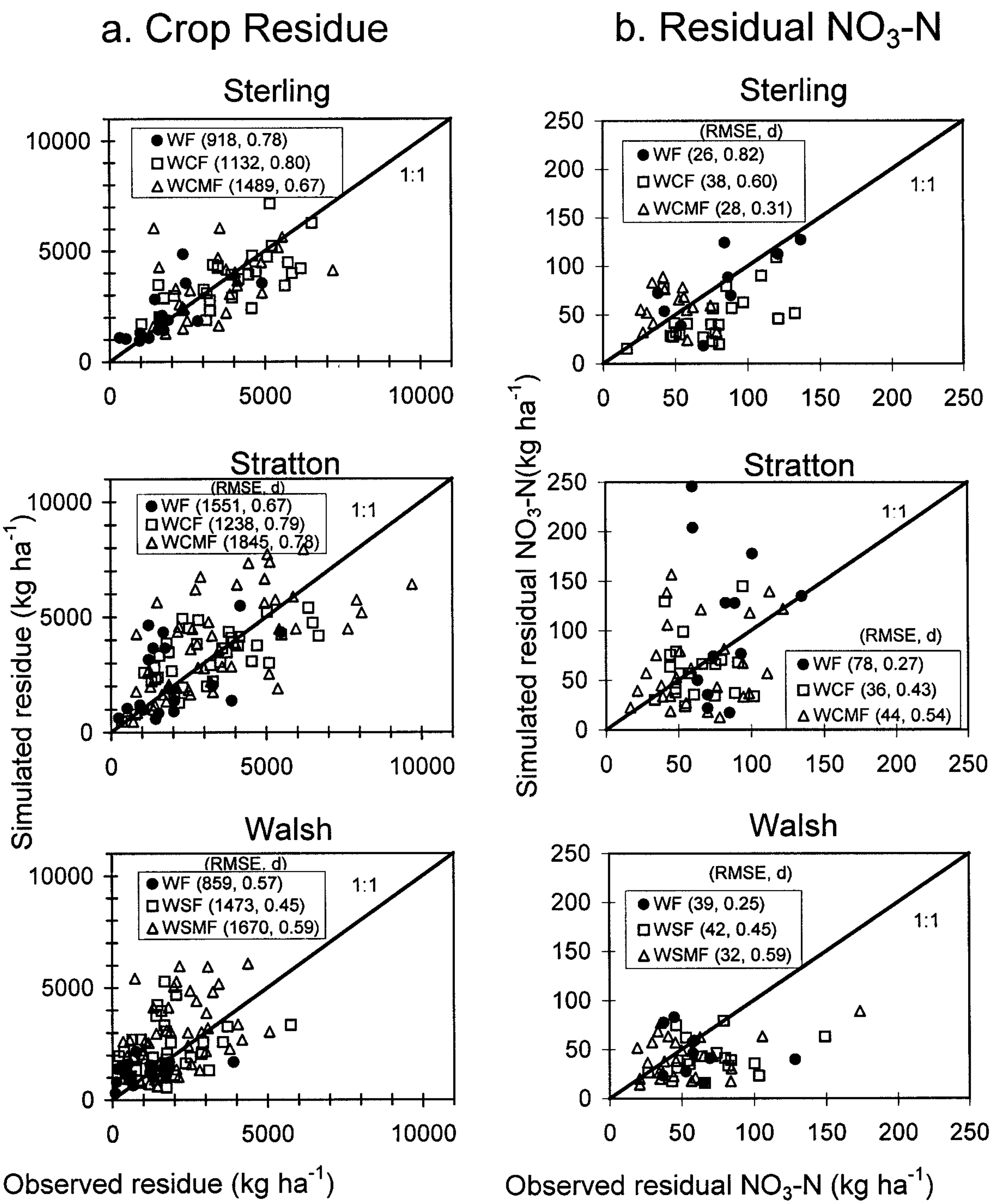

Fig. 7. Simulated (a) dry mass surface crop residue and (b) total soil profile residual NO_3–N against the observed values at three locations (summit position) in eastern Colorado for three rotations {wheat–fallow (WF): 1988–1999; wheat–corn (or sorghum)–fallow [WC(S)F]: 1988–1999; wheat–corn (or sorghum)–millet–fallow [WC(S)MF]: 1988–1993}. Values in parentheses are root mean square error (RMSE) (kg ha^{-1}) and *d* values.

between them are assumed to be reasonable since they presumably contain the same error (i.e., errors cancel out when considering differences in WUE). Interestingly, the GPFARM-simulated WUEs showed relative differences similar to those in the observed WUEs, particularly in matching the trend. Overpredictions in corn and proso millet yields in the WC(S)F and WC(S)MF systems may have contributed to the discrepancies between estimated and simulated WUE. Nevertheless, both the simulated and observed WUE showed a relatively large incremental increase when going from the WF to the WC(S)F rotation but showed very little (and even negative) incremental increase when shifting from WC(S)F to the WC(S)MF rotation. These results were consistent with the findings of Peterson et al. (1993) and Farahani et al. (1998) at the same locations. Again,

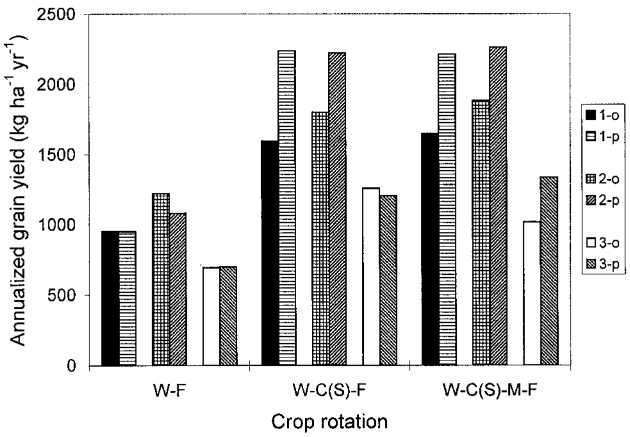

Fig. 8. Annualized grain yields (dry mass) at three locations (1 = Sterling, 2 = Stratton, and 3 = Walsh) in three crop rotations for the period 1989 through 1993. The legend indicates location and either observed (o) or predicted (p) value (e.g., 1-p = Sterling predicted). WF, wheat–fallow (rotation); WC(S)F, wheat–corn (or sorghum)–fallow (rotation); and WC(S)MF, wheat–corn (or sorghum)–millet–fallow (rotation).

GPFARM was able to simulate the observed pattern of WUE across the three locations: WUE was highest at Stratton and least at Walsh.

GPFARM captured the trends in average crop residues very well (Fig. 10). The observed data on crop residues are averages of preplant and preharvest residue measurements taken during the 1989 to 1993 period. The increased amount of crop residue maintained at the soil surface with the WC(S)F and WC(S)MF rotations compared with WF is another benefit of increased cropping intensity with no-till management. Crop residues on the soil surface protect the soil from erosion, reduce soil evaporation, and contribute organic matter to the soil when they decay.

DISCUSSION

The general purpose of GPFARM is to serve as a whole farm/ranch DSS for strategic (long-term) planning across the Great Plains, including production, economic and environmental impact analysis, and site-specific database generation, from which alternative agricultural management systems can be tested and compared. In crop production, the analysis of the economic viability of alternative management systems using a DSS such as GPFARM depends on accurate simulation of economic yield over a wide range of environmental and management conditions. Therefore, the model not only needs to accurately simulate the crop environment (e.g., weather, soil water balance, amount of surface residue, etc.) but also accurately simulate crop growth and yield. The results presented in this paper provide the limits of accuracy (RE, RMSE, and d values) within which GPFARM may be used to gauge performance of crops, as influenced by environmental variables, in alternative cropping systems under dryland, water-stressed conditions. Whereas the RE is an arithmetic average over the duration of data (i.e., shows long-term bias), the RMSE and d values indicate the average event-by-event (short-term) prediction errors. The acceptable limits of accuracy depend on the user needs, the type of management practices being compared and economic value of the differences between them, and whether the interest is in long-term or short-term differences (or both). GPFARM was shown to be better for evaluating long-term average differences or trends than for short-term comparisons.

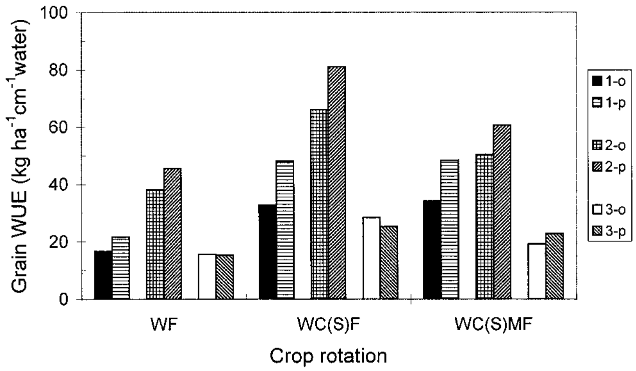

Fig. 9. Grain water use efficiency (WUE) at three locations (1 = Sterling, 2 = Stratton, and 3 = Walsh) in three crop rotations for the period 1989 through 1993. The legend indicates location and either observed (o) or predicted (p) value (e.g., 1-p = Sterling predicted). WF, wheat–fallow (rotation); WC(S)F, wheat–corn (or sorghum)–fallow (rotation); and WC(S)MF, wheat–corn (or sorghum)–millet–fallow (rotation).

The EPIC-based crop growth model in GPFARM appears to be more appropriate in estimating long-term average crop yields or trends in yields rather than simulating year-to-year variability in crop yields in eastern Colorado. This agrees with the findings of other investigators who tested the EPIC crop growth model. Kiniry et al. (1995) observed that EPIC can give reasonable mean yield simulations for the major crops and forages in the northern Great Plains but was unable to adequately simulate yield in some low-yielding years. This is consistent with the overpredictions of corn and proso millet grain yields that we observed in low-yielding dryland conditions in eastern Colorado. Similar to our findings, Kiniry et al. (1995) also observed EPIC's inability to simulate year-to-year variability in yield. Jara and Stockle (1999) found that EPIC performed poorly in

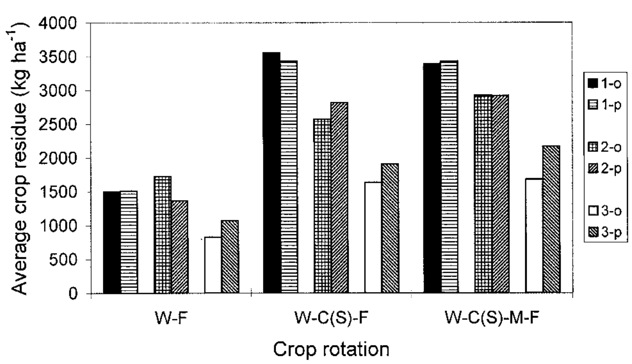

Fig. 10. Average crop residue on the soil surface at three locations (1 = Sterling, 2 = Stratton, and 3 = Walsh) in three crop rotations for the period 1989 through 1993. The legend indicates location and either observed (o) or predicted (p) value (e.g., 1-p = Sterling predicted). WF, wheat–fallow (rotation); WC(S)F, wheat–corn (or sorghum)–fallow (rotation); and WC(S)MF, wheat–corn (or sorghum)–millet–fallow (rotation).

simulating corn water uptake under water stress. Cabelguenne et al. (1999) observed that EPIC overestimated vegetative biomass and grain production, especially under conditions of pronounced water stress. All these suggest that the dryland conditions in eastern Colorado, which are characterized by periods of extreme water and temperature stresses, may be outside the validity domain of the EPIC crop model. However, the use of a generic crop model in GPFARM, as opposed to having several crop-specific (more process detail) models, greatly simplifies parameterization for many different crops grown in the Great Plains. The unavailability of within-season growth data made verification of the crop growth model difficult, and calibration was based only on final grain yield and biomass data. This may have been a greater limitation with corn, proso millet, and sorghum, which are exposed to more of the water and high-temperature stresses in summer compared with winter wheat. Therefore, more rigorous testing and improvement of the EPIC-based crop model in GPFARM must be done under dryland conditions in eastern Colorado using detailed observations of biomass, leaf area index, phenology, HI, and grain yield for various crops grown in the area. The correct simulation of crop response to extreme water stresses prevalent in eastern Colorado must be ascertained before making attempts at simulating more complex factors such as weed competition, freeze damage, and erratic emergence. Recalibration of the crop parameters and modification of the stress functions in the crop model may be required.

Also, recent enhancements made to the generic EPIC crop model may be integrated into GPFARM. For example, Cavero et al. (2000) and Cabelguenne et al. (1999) incorporated enhancements in their EPICPhase version of the crop model to improve simulations of root water uptake, leaf area, biomass accumulation, and water stress response. Xie et al. (2001) showed that ALMANAC (Kiniry et al., 1992), which is also based on the EPIC crop model, performed as well as or even better than the crop-specific CERES-Maize (Jones and Kiniry, 1986) and SORKAM sorghum model (Rosenthal et al., 1989) in simulating single-year corn and sorghum grain yields under water-limiting conditions.

As a temporary fix, calibration of the crop growth model separately for different levels of water stress may improve the predictions within those stress levels. For example, we found that the calibration of LAI_{max} and HI under Sterling toeslope conditions gave better predictions of grain yields at the toeslope than at the Sterling summit position with higher water stress levels. The RE values for corn grain yields were 28.7% (WCF) and 4.8% (WCMF) at the toeslope vs. 52.1% (WCF) and 47.6% (WCMF) at the summit position. However, this practice violates the scientific method and only compensates for the inadequacies of the water stress functions in the model. Thus, we favor the aforementioned testing and improvement of the crop model that should be based on sound theories of the mechanisms of crop response to varying levels of water stress.

One's perception of model accuracy is highly dependent on scale. For example, Rasse et al. (2000) pointed out that experimental plot data are not as buffered against pest damage and individual management errors as averaged county grain yields. They cited two contrasting studies: one that reported accurate grain yield simulation using averaged county grain yields (Kiniry et al., 1997) and another that reported poor simulation of year-to-year grain yield variability using research plot data (Otegui et al., 1996). GPFARM was developed to operate at the field scale, and both inputs (e.g., soil physical properties) and outputs (e.g., grain yield) of the model represent conditions averaged over an entire field. On the other hand, the observed data used to evaluate the model were taken at the plot scale, which is subject to greater variability. This may partly explain the apparent inadequacy of the model in simulating the observed variability in grain yields.

The GPFARM simulations of total soil water content in the profile were comparable in accuracy to those of RZWQM, which simulates the soil water balance with greater process detail. The indices of agreement (d) between GPFARM and observed total soil profile water content ranged from 0.64 to 0.81 across the three sites. In comparison, Wu et al. (1999) reported lower d values (0.54–0.59) for total water content simulations of RZWQM during two seasons in a sandy soil near Princeton, MN. The average RMSE in GPFARM-simulated total volumetric water content of the soil profile ranged from 0.029 to 0.052 across the three locations over multiple years. For RZWQM, Ma et al. (2002) reported RMSE values for total soil profile water content of 0.023 and 0.027 $cm^3\ cm^{-3}$ for a single irrigated corn and soybean season (calibration), respectively. Naturally, RMSE values tend to be greater when calculated over multiple years of validation than when calculated for just one season of model calibration. The errors in soil water content simulations are possibly well within the range of spatial variability, considering that only two point measurements were taken per treatment (1500-m^2 average plot area per treatment). Soil spatial variability—another cause of spatial yield variability because of its influence on soil water availability, fertility, and root distribution—was not considered in the simulations as only one soil profile was used for each location. The simulations of soil water content may also be improved by more accurate representation of rainfall intensities instead of assuming 2-h duration for all storms. The simulation of root distribution and root water uptake also need further investigation. Furthermore, in this study, all of the cropping systems were under no-till management, with significant amounts of crop residue on the soil surface. Previous studies have shown that rainfall interception (and subsequent water absorption) by residue can be a significant portion of total rainfall depth (Mohamoud and Ewing, 1990; Savabi and Stott, 1994). Thus, interception by crop residue can significantly reduce infiltration, especially during low-intensity rainfall events occurring over dry crop residues. The simulation of rainfall interception by residues in GPFARM would likely improve the soil water content simulations. The simulations of total soil residual NO_3–N look encouraging, but further evaluations of crop nitrate N uptake and N dynamics are needed.

Although the current version of GPFARM (v. 2.01) was found to be less suited for year-to-year grain yield prediction under dryland conditions in eastern Colorado, it has potential as a heuristic tool for studying long-term interactions between environment and crop management system. This was demonstrated by the ability of GPFARM to simulate long-term mean trends in annualized yields, WUE, and crop residue levels for three cropping systems across three locations in eastern Colorado. The user-friendly graphical user interface made it exceptionally easy for us to set up numerous combinations of management scenarios (planting, fertilizer application, harvest, and crop sequence), resources (soil, crop residues, initial soil water content, and soil residual NO_3–N), and climate. Improvements in the accuracy of the model's grain yield simulations—accounting for adverse factors such as extreme water stress, weeds, freeze damage, and hard soil surface conditions—will eventually fulfill the goal for GPFARM to serve as a DSS for production and economic impact analysis of alternative management systems. However, this will entail considerable efforts in further testing, model improvement, and field verification.

On a more practical note, it is important to constantly consider the needs and requirements of the intended users of GPFARM—producers and agricultural consultants in the Great Plains. Producers who have acquired practical management knowledge feel less need to adopt computer technology (e.g., DSS) whereas those without the practical skills view this as a means to compensate (Ascough et al., 1999). In the case of producer cooperators involved in testing GPFARM, it appears that their interest has shifted from a DSS for strategic (long-term) planning to one for tactical (real-time) decision-making (McMaster et al., 2003). And when producers are asked how accurate crop yield predictions must be, the typical response is "within about 10%." McMaster et al. (2003) note that the current state-of-the-art of crop growth simulation models is that, even with accurate determination of inputs and crop/variety parameters and no biotic stresses (i.e., weeds, pests, and diseases), it is extremely difficult to achieve the 10% requirement because of our inability to address all of the spatial and temporal variability inherent in the soils, climate, and management practices as well as deficiencies in our understanding of crop responses to stresses. These issues have raised the question of whether or not a generic crop growth model is adequate for GPFARM and highlight the great challenges in taking agricultural systems simulation technology outside of the research community and making it available to producers and agricultural advisors.

ACKNOWLEDGMENTS

The authors recognize the concerted effort of the GPFARM Team at the USDA-ARS Great Plains Systems Research Unit at Fort Collins, CO, in striving to develop a high quality, user-friendly decision support tool for farmers and ranchers in the northern Great Plains. Pat Bartling, Nam Ho, and Bruce Vandenberg provided programming and technical assistance to make the GPFARM simulations possible. Appreciation is also extended to Lucretia Sherrod, USDA-ARS technician, for providing all of the processed experimental data from the Sustainable Dryland Agroecosystem Management project; Dr. Gregory S. McMaster for giving suggestions in crop parameterization; and Dr. James C. Ascough II, Dr. Liwang Ma, and Dr. Dwayne G. Westfall for reviewing the manuscript.

REFERENCES

Ahuja, L.R., D.K. Cassel, R.R. Bruce, and B.B. Barnes. 1989. Evaluation of spatial distribution of hydraulic conductivity using effective porosity data. Soil Sci. 148:404–411.

Ahuja, L.R., W.J. Rawls, D.R. Neilsen, and R.D. Williams. 1999. Determining soil hydraulic properties and their field variability from simpler measurements. p. 1207–1233. *In* Agricultural drainage. Agron. Monogr. 38. ASA, CSSA, and SSSA, Madison, WI.

Ahuja, L.R., K.W. Rojas, J.D. Hanson, M.J. Shaffer, and L. Ma (ed.). 2000. Root Zone Water Quality Model: Modeling management effects on water quality and crop production. Water Resour. Publ., Highlands Ranch, CO.

Arnold, J.G., M.A. Weltz, E.E. Alberts, and D.C. Flanagan. 1995. Plant growth component. p. 8.1–8.41. *In* D.C. Flanagan, M.A. Nearing, and J.M. Laflen (ed.) USDA water erosion prediction project: Hillslope profile and watershed model documentation. NSERL Rep. 10. ARS Natl. Soil Erosion Res. Lab., West Lafayette, IN.

Ascough, J.C., II, D.L. Hoag, W.M. Frasier, and G.S. McMaster. 1999. Computer use in agriculture: An analysis of Great Plains producers. Comput. Electron. Agric. 23:189–204.

Ascough, J.C., II, M.J. Shaffer, D.L. Hoag, G.S. McMaster, G.H. Dunn, L.R. Ahuja, and M.A. Weltz. 2002. GPFARM: An integrated decision support system for sustainable Great Plains agriculture. p. 951–960. *In* D.E. Stott, R.H. Mohtar, and G.C. Steinhardt (ed.) Sustaining the global farm—local action for land leadership. Proc. Int. Soil Conserv. Org. (ISCO) Conf., 10th, Purdue Univ., West Lafayette, IN. 24–29 May 1999. Purdue Univ., West Lafayette, IN.

Brooks, R.H., and A.T. Corey. 1964. Hydraulic properties of porous media. Hydrology Paper 3. Colorado State Univ., Fort Collins.

Cabelguenne, M., P. Debaeke, and A. Bouniols. 1999. EPICphase, a version of the EPIC model simulating the effects of water and nitrogen stress on biomass and yield, taking account of developmental stages: Validation on maize, sunflower, sorghum, soybean and winter wheat. Agric. Syst. 60:175–196.

Campbell, G.S. 1974. A simple method for determining unsaturated conductivity from moisture retention data. Soil Sci. 117:311–314.

Cavero, J., I. Farre, P. Debaeke, and J.M. Faci. 2000. Simulation of maize yield under water stress with the EPICphase and CROPWAT models. Agron. J. 92:679–690.

Darcy, H. 1856. Les fontaines publiques de la Ville de Dijon. Dalmont, Paris.

Deer-Ascough, L.A., G.S. McMaster, J.C. Ascough II, and G.A. Peterson. 1998. Application of generic crop growth model technology to Great Plains conservation systems. ASAE Paper 98–2133. ASAE, St. Joseph, MI.

Farahani, H.J., and L.R. Ahuja. 1996. Evapotranspiration modeling of partial canopy/residue-covered fields. Trans. ASAE 39(6): 2051–2064.

Farahani, H.J., G.A. Peterson, D.G. Westfall, L.A. Sherrod, and L.R. Ahuja. 1998. Soil water storage in dryland cropping systems: The significance of cropping intensification. Soil Sci. Soc. Am. J. 62: 984–991.

Frasier, W.M., D.L. Hoag, and J. Ascough II. 1997. Computer use in agriculture: Opportunities for farm advisors. J. Am. Soc. Farm Managers Rural Appraisers 61:50–54.

Green, W.H., and G.A. Ampt. 1911. Studies on soil physics: I. Flow of air and water through soils. J. Agric. Sci. 4:1–24.

Jara, J., and C.O. Stockle. 1999. Simulation of water uptake in maize, using different levels of process detail. Agron. J. 91:256–265.

Jones, C.A., and J.R. Kiniry. (ed.) 1986. CERES-Maize: A simulation model of maize growth and development. Texas A&M Univ. Press, College Station.

Kiniry, J.R., D.J. Major, R.C. Izaurralde, J.R. Williams, P.W. Gassman, M. Morrison, R. Bergentine, and R.P. Zentner. 1995. EPIC model parameters for cereal, oilseed, and forage crops in the northern Great Plains region. Can. J. Plant Sci. 75:679–688.

Kiniry, J.R., J.R. Williams, P.W. Gassman, and P. Debaeke. 1992. A general, process-oriented model for two competing plant species. Trans. ASAE 35:801–810.

Kiniry, J.R., J.R. Williams, R.L. Vanderlip, J.D. Atwood, D.C. Reicosky, J. Mulliken, W.J. Cox, H.J. Mascagni, Jr., S.E. Hollinger, and W.J. Wiebold. 1997. Evaluation of two maize models for nine U.S. locations. Agron. J. 89:421–426.

Ma, L., D.C. Nielsen, L.R. Ahuja, J.R. Kiniry, J.D. Hanson, and G. Hoogenboom. 2002. An evaluation of RZWQM, CROPGRO, and CERES-Maize for responses to water stress in the Central Great Plains of the U.S. p. 119–148. *In* L.R. Ahuja et al. (ed.) Agricultural system models in field research and technology transfer. Lewis Publ., Boca Raton, FL.

Martin, S.M., M.A. Nearing, and R.R. Bruce. 1993. An evaluation of the EPIC model for soybeans grown in Southern Piedmont soils. Trans. ASAE 36(5):1327–1331.

McMaster, G.S., J.C. Ascough II, G.H. Dunn, M.A. Weltz, M.J. Shaffer, D. Palic, B.C. Vandenberg, P.N.S. Bartling, D. Edmunds, D.L. Hoag, and L.R. Ahuja. 2003. Application and testing of GPFARM: A farm and ranch decision support system for evaluating economic and environmental sustainability of agricultural enterprises. Acta Hortic. 593:171–177.

Mohamoud, Y.M., and L.K. Ewing. 1990. Rainfall interception by corn and soybean residue. Trans. ASAE 33(2):507–511.

Monteith, J.L. 1977. Climate and the efficiency of crop production in Britain. Philos. Trans. R. Soc. London, Ser. B 281:277–329.

Moulin, A.P., and H.J. Beckie. 1993. Evaluation of the CERES and EPIC models for predicting spring wheat grain yield over time. Can. J. Plant Sci. 73(3):713–719.

Otegui, M.E., R.A. Ruiz, and D. Petruzzi. 1996. Modeling hybrid and sowing date effects on potential grain yield of maize in a humid temperate region. Field Crops Res. 47:167–174.

Peterson, G.A., D.G. Westfall, and C.V. Cole. 1993. Agroecosystem approach to soil and crop management research. Soil Sci. Soc. Am. J. 57:1354–1360.

Peterson, G.A., D.G. Westfall, F.B. Peairs, L. Sherrod, D. Poss, W. Gangloff, K. Larson, D.L. Thompson, L.R. Ahuja, M.D. Koch, and C.B. Walker. 2000. Sustainable dryland agroecosystem management. Agric. Exp. Stn. Tech. Bull. TB00–3. Colorado State Univ., Fort Collins.

Peterson, G.A., D.G. Westfall, L. Sherrod, R. Kolberg, and D. Poss. 1995. Sustainable dryland agroecosystem management. Agric. Exp. Stn. Tech. Bull. TB95–1. Colorado State Univ., Fort Collins.

Rasse, D.P., J.T. Ritchie, W.W. Wilhelm, J. Wei, and E.C. Martin. 2000. Simulating inbred-maize yields with CERES-IM. Agron. J. 92:672–678.

Rawls, W.J. and D.L. Brakensiek. 1985. Prediction of soil water properties for hydrologic modeling. p. 293–299. *In* Proc. Watershed Manage. in the Eighties, Denver, CO. 30 April–1 May 1985. ASCE, New York.

Rosenthal, W.D., R.L. Vanderlip, B.S. Jackson, and G.F. Arkin. 1989. SORKAM: A grain sorghum crop growth model. Computer Software Documentation Ser. MP 1669. Texas Agric. Exp. Stn., College Station.

Savabi, M.R., and D.E. Stott. 1994. Plant residue impact on rainfall interception. Trans. ASAE 37(4):1093–1098.

Shaffer, M.J., P.N.S. Bartling, and J.C. Ascough II. 2000. Object-oriented simulation of integrated whole farms: GPFARM framework. Comput. Electron. Agric. 28:29–49.

Shaffer, M.J., A.D. Halvorson, and F.J. Pierce. 1991. Nitrate Leaching and Economic Analysis Package (NLEAP): Model description and application. p. 285–322. *In* R. Follett et al. (ed.) Managing nitrogen for groundwater quality and farm profitability. SSSA, Madison, WI.

Shaffer, M.J., K. Lasnik, X. Ou, and R. Flynn. 2001. NLEAP internet tools for estimating NO_3–N leaching and N_2O emissions. p. 403–426. *In* M.J. Shaffer et al. (ed.) Modeling carbon and nitrogen dynamics for soil management. CRC Press, Boca Raton, FL.

Steiner, J.L., J.R. Williams, and O.R. Jones. 1987. Evaluation of the EPIC simulation model using a dryland wheat–sorghum–fallow crop rotation. Agron. J. 79:732–738.

Williams, J.R., C.A. Jones, and P.T. Dyke. 1984. A modeling approach to determining the relationship between erosion and soil productivity. Trans. ASAE 27(1):129–144.

Williams, J.R., C.A. Jones, J.R. Kiniry, and D.A. Spanel. 1989. The EPIC crop growth model. Trans. ASAE 32:497–511.

Willmott, C.J. 1981. On the validation of models. Phys. Geogr. 2: 184–194.

Willmott, C.J., and D.E. Wicks. 1980. An empirical method for the spatial interpolation of monthly precipitation within California. Phys. Geogr. 1:59–73.

Wu, L., W. Chen, J.M. Baker, and J.A. Lamb. 1999. Evaluation of the Root Zone Water Quality Model using field-measured data from a sandy soil. Agron. J. 91:177–182.

Xie, Y., J.R. Kiniry, V. Nedbalek, and W.D. Rosenthal. 2001. Maize and sorghum simulations with CERES-Maize, SORKAM, and ALMANAC under water-limiting conditions. Agron. J. 93:1148–1155.